

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5

(i) APPLICANT: Dalla-Favera, Riccardo

(ii) TITLE OF INVENTION: IDENTIFICATION OF GENES ALTERED IN
MULTIPLE MYELOMA

10

(iii) NUMBER OF SEQUENCES: 17

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Cooper & Dunham LLP

15

(B) STREET: 1185 Avenue of the Americas

(C) CITY: New York

(D) STATE: New York

(E) COUNTRY: U.S.A.

(F) ZIP: 10036

20

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

25

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE: 28-MAY-1996

30

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: White, John P.

(B) REGISTRATION NUMBER: 28,678

35

(C) REFERENCE/DOCKET NUMBER: 50995

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (212) 278-0400

(B) TELEFAX: (212) 391-0525

40

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 108 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Lys Leu Arg Gln Trp Leu Ile Asp Gln Ile Asp Ser Gly Lys Tyr Pro
1 5 10 15

Gly Leu Val Trp Glu Asn Glu Glu Lys Ser Ile Phe Arg Ile Pro Trp
20 25 30

Lys His Ala Gly Lys Gln Asp Tyr Asn Arg Glu Glu Asp Ala Ala Leu
35 40 45

Phe Lys Ala Trp Ala Leu Phe Lys Gly Lys Phe Arg Glu Gly Ile Asp
50 55 60

Lys Pro Asp Pro Pro Thr Trp Lys Thr Arg Leu Arg Cys Ala Leu Asn
65 70 75 80

Lys Ser Asn Asp Phe Glu Glu Leu Val Glu Arg Ser Gln Leu Asp Ile
85 90 95

Ser Asp Pro Tyr Lys Val Tyr Arg Ile Val Pro Glu
100 105

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 108 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

10 Lys Leu Arg Gln Trp Leu Ile Asp Gln Ile Asp Ser Gly Lys Tyr Pro
1 5 10 15

Gly Leu Val Trp Glu Asn Glu Glu Lys Ser Val Phe Arg Ile Pro Trp
20 25 30

15 Lys His Ala Gly Lys Gln Asp Tyr Asn Arg Glu Glu Asp Ala Ala Leu
35 40 45

Phe Lys Ala Trp Ala Leu Phe Lys Gly Lys Phe Arg Glu Gly Ile Asp
20 50 55 60

Lys Pro Asp Pro Pro Thr Trp Lys Thr Arg Leu Arg Cys Ala Leu Asn
65 70 75 80

25 Lys Ser Asn Asp Phe Glu Glu Leu Val Glu Arg Ser Gln Leu Asp Ile
85 90 95

Ser Asp Pro Tyr Lys Val Tyr Arg Ile Val Pro Glu
100 105

30

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 108 amino acids

35

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Arg Met Arg Pro Trp Leu Glu Met Gln Ile Asn Ser Asn Gln Ile Pro
 1 5 10 15

Gly Leu Ile Trp Ile Asn Lys Glu Glu Met Ile Phe Gln Ile Pro Trp
 20 25 30

Lys His Ala Ala Lys His Gly Trp Asp Ile Asn Lys Asp Ala Cys Leu
 35 40 45

Phe Arg Ser Trp Ala Ile His Thr Gly Arg Tyr Lys Ala Gly Glu Lys
 50 55 60

Glu Pro Asp Pro Lys Thr Trp Lys Ala Asn Phe Arg Cys Ala Met Asn
 65 70 75 80

Ser Leu Pro Asp Ile Glu Glu Val Lys Asp Gln Lys Arg Asn Lys Gly
 85 90 95

Ser Ser Ala Val Arg Val Tyr Arg Met Leu Pro Pro
 100 105

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 108 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Arg Met Arg Pro Trp Leu Glu Glu Gln Ile Asn Ser Asn Thr Ile Pro
 1 5 10 15

Gly Leu Lys Trp Leu Asn Lys Glu Lys Lys Ile Phe Gln Ile Pro Trp

30

5

10

15

(2) INFORMATION FOR SEQ ID NO:5:

20

25

(ii) MOLECULE TYPE: peptide

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

35

40

Phe Lys Ala Trp Ala Val Phe Lys Gly Lys Phe Lys Glu Gly Asp Lys

50

55

60

Ala Glu Pro Ala Thr Trp Lys Thr Arg Leu Arg Cys Ala Leu Asn Lys
65 70 75 80

5

Ser Pro Asp Phe Glu Glu Val Thr Asp Arg Ser Gln Leu Asp Ile Ser
85 90 95

Glu Pro Tyr Lys Val Tyr Arg Ile Val Pro Glu
100 105

10

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 107 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: peptide

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Lys Leu Arg Asn Trp Val Val Glu Gln Val Glu Ser Gly Gln Phe Pro
1 5 10 15

30

Gly Val Cys Trp Asp Asp Thr Ala Lys Thr Met Phe Arg Ile Pro Trp
20 25 30

Lys His Ala Gly Lys Gln Asp Phe Arg Glu Asp Gln Asp Ala Ala Phe
35 40 45

35

Phe Lys Ala Trp Ala Ile Phe Lys Gly Lys Tyr Lys Glu Gly Asp Thr
50 55 60

40

Gly Gly Pro Ala Val Trp Lys Thr Arg Leu Arg Cys Ala Leu Asn Lys
65 70 75 80

Ser Ser Glu Phe Lys Glu Val Pro Glu Arg Gly Arg Met Asp Val Ala

85

90

95

Glu Pro Tyr Lys Val Tyr Gln Leu Leu Pro Pro
100 105

5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 107 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

20

Arg Ile Leu Pro Trp Leu Val Ser Gln Leu Asp Leu Gly Gln Leu Glu
1 5 10 15

25

Gly Val Ala Trp Val Asn Lys Ser Arg Thr Arg Phe Arg Ile Pro Trp
20 25 30

Lys His Gly Leu Arg Gln Asp Ala Gln Gln Glu Asp Phe Gly Ile Phe
35 40 45

30

Gln Ala Trp Ala Glu Ala Thr Gly Ala Tyr Val Pro Gly Arg Asp Lys
50 55 60

Pro Asp Leu Pro Thr Trp Lys Arg Asn Phe Arg Ser Ser Ala Leu Asn
65 70 75 80

35

Arg Lys Glu Gly Leu Arg Leu Ala Glu Asp Arg Ser Lys Asp Pro His
85 90 95

40

Asp Pro His Lys Ile Tyr Glu Phe Val Asn Ser
100 105

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 95 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Lys Arg Leu Cys Gln Ser Thr Ile Tyr Trp Asp Gly Pro Leu Ala Leu
1 5 10 15

Cys Asn Asp Arg Pro Asn Lys Leu Glu Arg Asp Gln Thr Cys Lys Leu
20 25 30

Phe Asp Thr Gln Gln Phe Leu Ser Glu Leu Gln Ala Phe Ala His His
35 40 45

Gly Arg Ser Leu Pro Arg Phe Gln Val Thr Leu Cys Phe Gly Glu Glu
50 55 60

Phe Pro Asp Pro Gln Arg Gln Arg Lys Leu Ile Thr Ala His Val Glu
65 70 75 80

Pro Leu Leu Ala Arg Gln Leu Tyr Tyr Phe Ala Gln Gln Asn Ser
85 90 95

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 95 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

Lys Arg Leu Cys Gln Ser Arg Ile Tyr Trp Asp Gly Pro Leu Ala Leu
1 5 10 15

Cys Ser Asp Arg Pro Asn Lys Leu Glu Arg Asp Gln Thr Cys Lys Leu
20 25 30

10 Phe Asp Thr Gln Gln Phe Leu Ser Glu Leu Gln Val Phe Ala His His
 35 40 45

Gly Arg Pro Ala Pro Arg Phe Gln Val Thr Leu Cys Phe Gly Glu Glu
50 55 60

15
Phe Pro Asp Pro Gln Arg Gln Arg Lys Leu Ile Thr Ala His Val Glu
65 70 75 80

Pro Leu Leu Ala Arg Gln Leu Tyr Tyr Phe Ala Gln Gln Asn Thr
20 85 90 95

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 96 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: peptide

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Lys Arg Leu Cys Gln Gly Arg Val Phe Cys Ser Gly Asn Ala Val Val
1 5 10 15

[illegible]

Gln Gly Arg Leu Pro Asp Gly Arg Val Val Leu Cys Phe Gly Glu Glu
50 55 60

Phe Pro Asp Met Ala Pro Leu Arg Ser Lys Leu Ile Leu Val Gln Ile
65 70 75 80

10 Glu Gln Leu Tyr Val Arg Gln Leu Ala Glu Glu Ala Gly Lys Ser Cys
 85 90 95

15

(A) LENGTH: 96 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Gln Arg Leu Cys Pro Ile Pro Ile Ser Trp Asn Ala Pro Gln Ala Pro
30 1 5 10 15

Pro Gly Pro Gly Pro His Leu Leu Pro Ser Asn Glu Cys Val Glu Leu
20 25 30

35 Phe Arg Thr Ala Tyr Phe Cys Arg Asp Leu Val Arg Tyr Phe Gln Gly
 35 40 45

Leu Gly Pro Pro Pro Lys Phe Gln Val Thr Leu Asn Phe Trp Glu Glu
50 55 60

40
Ser His Gly Ser Ser His Thr Pro Gln Asn Leu Ile Thr Val Lys Met
65 70 75 80

Glu Gln Ala Phe Ala Arg Tyr Leu Leu Glu Gln Thr Pro Glu Gln Gln
85 90 95

5 (2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 100 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gln Arg Leu Gly His Cys His Thr Tyr Trp Ala Val Ser Glu Glu Leu
1 5 10 15
Leu Pro Asn Ser Gly His Gly Pro Asp Gly Glu Val Pro Lys Asp Lys
20 25 30
Glu Gly Gly Val Phe Asp Leu Gly Pro Phe Ile Val Asp Leu Ile Thr
35 40 45
Phe Thr Glu Gly Ser Gly Arg Ser Pro Arg Tyr Ala Trp Leu Phe Cys
50 55 60
Val Gly Glu Ser Trp Pro Gln Asp Gln Pro Trp Thr Lys Arg Leu Val
65 70 75 80
Met Val Lys Val Val Pro Thr Cys Leu Arg Ala Leu Val Glu Met Ala
85 90 95
Arg Val Gly Gly
100

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5176 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 217..1569

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GCCTGACCAA CATGGTAAAA CCCCATCTCT GCTAAACTA CAAAAAATTA GCTGGATGTG
60

GTGGCAGGGA ACCTGTCATC CCAGCTAGTT GGGAGACTGA GGCAGGAGAA TCGCTCGATC
120

TTGGGACCCA CCGCTGCCCT CAGCTCCGAG TCCAGGGCGA GTGCAGAGCA CAGCGGGCGG
180

AGGACCCCGG GCGCGGGCGC GGACGGCACG CGGGGC ATG AAC CTG GAG GGC GGC
234

Met Asn Leu Glu Gly Gly
1 5

GGC CGA GGC GGA GAG TTC GGC ATG AGC GCG GTG AGC TGC GGC AAC GGG
282

Gly Arg Gly Gly Glu Phe Gly Met Ser Ala Val Ser Cys Gly Asn Gly
10 15 20

AAG CTC CGC CAG TGG CTG ATC GAC CAG ATC GAC AGC GGC AAG TAC CCC
330

Lys Leu Arg Gln Trp Leu Ile Asp Gln Ile Asp Ser Gly Lys Tyr Pro
25 30 35

GGG CTG GTG TGG GAG AAC GAG GAG AAG AGC ATC TTC CGC ATC CCC TGG
378

AAG CAC GCG GGC AAG CAG GAC TAC AAC CGC GAG GAG GAC GCC GCG CTC
426

TTC AAG GCT TGG GCA CTG TTT AAA GGA AAG TTC CGA GAA GGC ATC GAC
474

AAG CCG GAC CCT CCC ACC TGG AAG ACG CGC CTG CGG TGC GCT TTG AAC
522

AAG AGC AAT GAC TTT GAG GAA CTG GTT GAG CGG AGC CAG CTG GAC ATC
570

TCA GAC CCG TAC AAA GTG TAC AGG ATT GTT CCT GAG GGA GCC AAA AAA
618

GGA GCC AAG CAG CTC ACC CTG GAG GAC CCG CAG ATG TCC ATG AGC CAC
666

CCC TAC ACC ATG ACA ACG CCT TAC CCT TCG CTC CCA GCC CAG CAG GTT
714

CAC AAC TAC ATG ATG CCA CCC CTC GAC CGA AGC TGG AGG GAC TAC GTC
762

His Asn Tyr Met Met Pro Pro Leu Asp Arg Ser Trp Arg Asp Tyr Val
170 175 180

CCG GAT CAG CCA CAC CCG GAA ATC CCG TAC CAA TGT CCC ATG ACG TTT
810

Pro Asp Gln Pro His Pro Glu Ile Pro Tyr Gln Cys Pro Met Thr Phe
185 190 195

5

GGA CCC CGC GGC CAC CAC TGG CAA GGC CCA GCT TGT GAA AAT GGT TGC
858

Gly Pro Arg Gly His His Trp Gln Gly Pro Ala Cys Glu Asn Gly Cys
200 205 210

10

CAG GTG ACA GGA ACC TTT TAT GCT TGT GCC CCA CCT GAG TCC CAG GCT
906

Gln Val Thr Gly Thr Phe Tyr Ala Cys Ala Pro Pro Glu Ser Gln Ala
215 220 225 230

15

CCC GGA GTC CCC ACA GAG CCA AGC ATA AGG TCT GCC GAA GCC TTG GCG
954

Pro Gly Val Pro Thr Glu Pro Ser Ile Arg Ser Ala Glu Ala Leu Ala
235 240 245

20

TTC TCA GAC TGC CGG CTG CAC ATC TGC CTG TAC TAC CGG GAA ATC CTC
1002

Phe Ser Asp Cys Arg Leu His Ile Cys Leu Tyr Tyr Arg Glu Ile Leu
250 255 260

25

GTG AAG GAG CTG ACC ACG TCC AGC CCC GAG GGC TGC CGG ATC TCC CAT
1050

Val Lys Glu Leu Thr Thr Ser Ser Pro Glu Gly Cys Arg Ile Ser His
265 270 275

30

GGA CAT ACG TAT GAC GCC AGC AAC CTG GAC CAG GTC CTG TTC CCC TAC
1098

Gly His Thr Tyr Asp Ala Ser Asn Leu Asp Gln Val Leu Phe Pro Tyr
280 285 290

35

CCA GAG GAC AAT GGC CAC AGG AAA AAC ATT GAG AAC CTG CTG AGC CAC
1146

Pro Glu Asp Asn Gly His Arg Lys Asn Ile Glu Asn Leu Leu Ser His
295 300 305 310

40

CTG GAG AGG GGC GTG GTC CTC TGG ATG GCC CCC GAC GGG CTC TAT GCG
1194

Leu Glu Arg Gly Val Val Leu Trp Met Ala Pro Asp Gly Leu Tyr Ala
315 320 325

AAA AGA CTG TGC CAG AGC ACG ATC TAC TGG GAC GGG CCC CTG GCG CTG
1242

Lys Arg Leu Cys Gln Ser Thr Ile Tyr Trp Asp Gly Pro Leu Ala Leu
330 335 340

TGC AAC GAC CGG CCC AAC AAA CTG GAG AGA GAC CAG ACC TGC AAG CTC
1290

Cys Asn Asp Arg Pro Asn Lys Leu Glu Arg Asp Gln Thr Cys Lys Leu
345 350 355

TTT GAC ACA CAG CAG TTC TTG TCA GAG CTG CAA GCG TTT GCT CAC CAC
1338

Phe Asp Thr Gln Gln Phe Leu Ser Glu Leu Gln Ala Phe Ala His His
360 365 370

GGC CGC TCC CTG CCA AGA TTC CAG GTG ACT CTA TGC TTT GGA GAG GAG
1386

Gly Arg Ser Leu Pro Arg Phe Gln Val Thr Leu Cys Phe Gly Glu Glu
375 380 385 390

TTT CCA GAC CCT CAG AGG CAA AGA AAG CTC ATC ACA GCT CAC GTA GAA
1434

Phe Pro Asp Pro Gln Arg Gln Arg Lys Leu Ile Thr Ala His Val Glu
395 400 405

CCT CTG CTA GCC AGA CAA CTA TAT TAT TTT GCT CAA CAA AAC AGT GGA
1482

Pro Leu Leu Ala Arg Gln Leu Tyr Tyr Phe Ala Gln Gln Asn Ser Gly
410 415 420

CAT TTC CTG AGG GGC TAC GAT TTA CCA GAA CAC ATC AGC AAT CCA GAA
1530

His Phe Leu Arg Gly Tyr Asp Leu Pro Glu His Ile Ser Asn Pro Glu
425 430 435

GAT TAC CAC AGA TCT ATC CGC CAT TCC TCT ATT CAA GAA TGAAAAATGT
1579

Asp Tyr His Arg Ser Ile Arg His Ser Ser Ile Gln Glu
440 445 450

GTCTTGCTCT GTCTCCCAGG CTGGAGTGCA GTGACACAAT CTCAGCTCAC TGTGACCTCC
1699

10 TGAGCCACTG CACCCACCCA AGACAAGTGA TTTTCATTGT AAATATTTGA CTTTAGTGAA
1819

15
AGCGGTTGAG GAGAATTGCG GCGAGACAAG CATGGAAAAT CAGTGACATC TGATTGGCAG
1939

CATTGATGAT CACTGTGAAA ATTGACCAAG TGATGTGTTT ACATTTACTG AAATGCGCTC
2059

ACACTGACTA GAGTGATGAC TGCTTGTAGG TATGTCTGTG CCATTTCTCA GGGAAGTAAG
2179

35 TTCTTTGTGGA AGACACTTGC TGAGTGAAGG AAATGAATCT TTGACTGAAG CCGTGCCTGT
2299

40 CACCCTCCTT CCCATTGGCT TTCTCTCCTT GGCCTTTCCT GGAAGCCAGT TAGTAAACTT
2419

CCTATTTTCT TGAGTCAAAA AACATGAGCG CTA CTCTTGG ATGGGACATT TTTGTCTGTC
2479

5 CTACAATCTA GTAATGTCTA AGTAATGGTT AAGTTTTCTT GTTTCTGCAT CTTTTTGACC
2539

CTCATTCTTT AGAGATGCTA AAATTCTTCG CATAAAGAAG AAGAAATTAA GGAACATAAA
2599

10 TCTTAATACT TGAAGTGTG CCCTTCTGTC CAAGTACTTA ACTATCTGTT CCCTTCCTCT
2659

15 GTGCCACGCT CCTCTGTTTG TTTGGCTGTC CAGCGATCAG CCATGGCGAC ACTAAAGGAG
2719

GAGGAGCCGG GGA CTCCAG GCTGGAGAGC ACTGCCAGGA CCCACCACTG GAAGCAGGAT
2779

20 GGAGCTGACT ACGGAACTGC AACTCAGTG GGCTGTTTCT GCTTATTTC TCTGTTCTAT
2839

GCTTCCTCGT GCCAATTATA GTTTGACAGG GCCTTAAAAT TACTTGGCTT TTTCCAAATG
2899

25 CTTCTATTTA TAGAAATCCC AAAGACCTCC ACTTGCTTAA GTATACCTAT CACTTACATT
2959

30 TTTGTGGTTT TGAGAAAGTA CAGCAGTAGA CTGGGGCGTC ACCTCCAGGC CGTTTCTCAT
3019

ACTACAGGAT ATTTACTATT ACTCCCAGGA TTCAGCAGAA GATTGCGTTA GCTCTCAAAT
3079

35 GTGTGTTTCT GCTTTTCTAA TGGATATTTT AAATTCATTC AACAAGCACC TAGTAAGTGC
3139

CTGCTGTATC CCTACATTAC ACAGTTCAGC CTTTATCAAG CTTAGTGAGC AGTGAGCACT
3199

40 GAAACATTAT TTTTAAATGT TTAAAAAGTT TCTAATATTA AAGTCAGAAT ATTAATACAA
3259

CCF090 "C0909090"

TTAATATTAA TATTAAC TAC AGAAAAGACA AACAGTAGAG AACAGCAAAA AAATAAAAAG
3319

5 GATCTCCTTT TTTCCCAGCC CAAATTCTCC TCTCTAAAAG TGTCCACAAG AAGGGGTGTT
3379

TATTCTTCCA ACACATTTCA CTTTTCTGTA AATATACATA AACTTAAAAA GAAAACCTCA
3439

10 TGGAGTCATC TTGCACACAC TTTTCATGCA GTGCTCTTTG TAGCTAAACA GTGAAGATTT
3499

15 ACCTCGTTCT GCTCAGAGGC CTTGCTGTGG AGCTCCACTG CCATGTACCC AGTAGGGTTT
3559

GACATTTTCAT TAGCCATGCA ACATGGATAT GTATTGGGCA GCAGACTGTG TTTCGTGAAC
3619

20 TGCAGTGATG TATACATCTT ATAGATGCAA AGTATTTTGG GGTATATTAT CCTAAGGGAA
3679

GATAAAGATG ATATTAAGAA CTGCTGTTTC ACGGGGCCCT TACCTGTGAC CCTCTTTGCT
3739

25 GAAGAATATT TAACCCACACA CAGCACTTCA AAGAAGCTGT CTTGGAAGTC TGTCTCAGGA
3799

30 GCACCCTGTC TTCTTAATTC TCCAAGCGGA TGCTCCATTT CAATTGCTTT GTGACTTCTT
3859

CTTCTTTGTT TTTTAAATA TTATGCTGCT TTAACAGTGG AGCTGAATTT TCTGGAAAAT
3919

35 GCTTCTTGGC TGGGGCCACT ACCTCCTTTC CTATCTTTAC ATCTATGTGT ATGTTGACTT
3979

TTTAAAATTC TGAGTGATCC AGGGTATGAC CTAGGGAATG AACTAGCTAT GGAAATAACT
4039

40 CAGGGTTAGG AATCCTAGCA CTTGTCTCAG GACTCTGAAA AGGAACGGCT TCCTCATTC
4099

TTGTCTTGAT AAAGTGGAAT TGGCAAATA GAATTTAGTT TGTACTCAGT GGACAGTGCT
4159

5 GTTGAAGATT TGAGGACTTG TTAAAGAGCA CTGGGTCATA TGGAAAAAAT GTATGTGTCT
4219

CCCCAGGTGC ATTTTCTTGG TTTATGTCTT GTTCTTGAGA TTTTGTATAT TTAGGAAAAC
4279

10 CTCAAGCAGT AATTAATATC TCCTGGAACA CTATAGAGAA CCAAGTGACC GACTCATTTA
4339

CAACTGAAAC CTAGGAAGCC CCTGAGTCCT GAGCGAAAAC AGGAGAGTTA GTCGCCCTAC
4399

15 AGAAAACCCA GCTAGACTAT TGGGTATGAA CTAAAAAGAG ACTGTGCCAT GGTGAGAAAA
4459

20 ATGTAAAATC CTACAGTGGA ATGAGCAGCC CTTACAGTGT TGTTACCACC AAGGGCAGGT
4519

AGGTATTAGT GTTTGAAAAA GCTGGTCTTT GAGCGAGGGC ATAAATACAG CTAGCCCCAG
4579

25 GGGTGGAACA ACTGTGGGAG TCTTGGGTAC TCGCACCTCT TGGCTTTGTT GATGCTCCGC
4639

CAGGAAGGCC ACTTGTGTGT GCGTGTCAGT TACTTTTTTA GTAACAATTC AGATCCAGTG
4699

30 TAAACTTCCG TTCATTGCTC TCCAGTCACA TGCCCCACT TCCCCACAGG TGAAAGTTTT
4759

35 TCTGAAGTGT TGGGATTGGT TAAGGTCTTT ATTTGTATTA CGTATCTCCC CAAGTCCTCT
4819

GTGGCCAGCT GCATCTGTCT GAATGGTGCG TGAAGGCTCT CAGACCTTAC ACACCATTTT
4879

40 GTAAGTTATG TTTTACATGC CCCGTTTTTG AGACTGATCT CGATGCAGGT GGATCTCCTT
4939

GAGATCCTGA TAGCCTGTTA CAGGAATGAA GTAAAGGTCA GTTTTTTTTGG TATTGATTTT
4999

5 CACAGCTTTG AGGAACATGC ATAAGAAATG TAGCTGAAGT AGAGGGGACG TGAGAGAAGG
5059

GCCAGGCCCGG CAGGCCAACC CTCCTCCAAT GGAAATTCCC GTGTTGCTTC AAAGTGAGAC
5119

10 AGATGGGACT TAACAGGCAA TGGGGTCCAC TTCCCCCTCT TCAGCATCCC CCGTACC
5176

(2) INFORMATION FOR SEQ ID NO:14:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 451 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

25

Met Asn Leu Glu Gly Gly Gly Arg Gly Gly Glu Phe Gly Met Ser Ala
1 5 10 15

Val Ser Cys Gly Asn Gly Lys Leu Arg Gln Trp Leu Ile Asp Gln Ile
20 25 30

30

Asp Ser Gly Lys Tyr Pro Gly Leu Val Trp Glu Asn Glu Glu Lys Ser
35 40 45

35

Ile Phe Arg Ile Pro Trp Lys His Ala Gly Lys Gln Asp Tyr Asn Arg
50 55 60

Glu Glu Asp Ala Ala Leu Phe Lys Ala Trp Ala Leu Phe Lys Gly Lys
65 70 75 80

40

Phe Arg Glu Gly Ile Asp Lys Pro Asp Pro Pro Thr Trp Lys Thr Arg
85 90 95

Pro Asp Gly Leu Tyr Ala Lys Arg Leu Cys Gln Ser Thr Ile Tyr Trp
325 330 335

5 Asp Gly Pro Leu Ala Leu Cys Asn Asp Arg Pro Asn Lys Leu Glu Arg
340 345 350

Asp Gln Thr Cys Lys Leu Phe Asp Thr Gln Gln Phe Leu Ser Glu Leu
355 360 365

10 Gln Ala Phe Ala His His Gly Arg Ser Leu Pro Arg Phe Gln Val Thr
370 375 380

15 Leu Cys Phe Gly Glu Glu Phe Pro Asp Pro Gln Arg Gln Arg Lys Leu
385 390 395 400

Ile Thr Ala His Val Glu Pro Leu Leu Ala Arg Gln Leu Tyr Tyr Phe
405 410 415

20 Ala Gln Gln Asn Ser Gly His Phe Leu Arg Gly Tyr Asp Leu Pro Glu
420 425 430

His Ile Ser Asn Pro Glu Asp Tyr His Arg Ser Ile Arg His Ser Ser
435 440 445

25 Ile Gln Glu
450

(2) INFORMATION FOR SEQ ID NO:15:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 152 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: DNA (genomic)

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TTTTCTCTAC AGTCACCTCC CTGTTTACCA AAGATAATCA CAATAAGTCC AGTTTACTTA
60

CAAAACAAGT TTAGTTATTA GAGGAACTA AACTTCAGG ATTCAGTCCA GATAATTTTT
120

AAAAACTCTA AAACAATGGA CAGGGCTAGA AT
152

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 152 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TGGGCTCGGC CTGGTGGGGC AGCCACAGCG GGACGCAGTA GTGAAAGTCC AGTTTACTTA
60

CAAAACAAGT TTAGTTATTA GAGGAACTA AACTTCAGG ATTCAGCAGG GCATGAGGAG
120

GCAGCTCCTC ACCCTCCCTT TCTCTTTTGT AC
152

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 152 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5

60

10

AGGCAGCTCC TCACCCTCCC TTTCTCTTTT GT

15

1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2